

Context Surrounding Processing Sites Is Crucial in Determining Cleavage Rate of a Subset of Processing Sites in HIV-1 Gag and Gag-Pro-Pol Polyprotein Precursors by Viral Protease*

Received for publication, January 4, 2012, and in revised form, February 9, 2012. Published, JBC Papers in Press, February 13, 2012, DOI 10.1074/jbc.M112.339374

Sook-Kyung Lee[‡], Marc Potempa[§], Madhavi Kolli[¶], Ayşegül Özen[¶], Celia A. Schiffer[¶], and Ronald Swanstrom^{‡§1}

From the [‡]Department of Biochemistry and Biophysics, and the University of North Carolina Center for AIDS Research,

[§]Department of Microbiology and Immunology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599 and

the [¶]Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, Massachusetts 01605

Background: Processing of HIV-1 Gag/Gag-Pro-Pol polyprotein by viral protease is an essential step to produce infectious virus particle.

Results: Processing of CA/SP1, SP1/NC, and SP2/p6 sites by HIV-1 protease are affected by context.

Conclusion: Complex substrate interactions beyond the active site of the enzyme define HIV-1 processing rate.

Significance: This study helps to understand role of context outside cleavage sites in HIV-1 processing.

Processing of the human immunodeficiency virus type 1 (HIV-1) Gag and Gag-Pro-Pol polyproteins by the HIV-1 protease (PR) is essential for the production of infectious particles. However, the determinants governing the rates of processing of these substrates are not clearly understood. We studied the effect of substrate context on processing by utilizing a novel protease assay in which a substrate containing HIV-1 matrix (MA) and the N-terminal domain of capsid (CA) is labeled with a FLA^{SH} (fluorescein arsenical hairpin) reagent. When the seven cleavage sites within the Gag and Gag-Pro-Pol polyproteins were placed at the MA/CA site, the rates of cleavage changed dramatically compared with that of the cognate sites in the natural context reported previously. The rate of processing was affected the most for three sites: CA/spacer peptide 1 (SP1) (~10-fold increase), SP1/nucleocapsid (NC) (~10–30-fold decrease), and SP2/p6 (~30-fold decrease). One of two multidrug-resistant (MDR) PR variants altered the pattern of processing rates significantly. Cleavage sites within the Pro-Pol region were cleaved in a context-independent manner, suggesting for these sites that the sequence itself was the determinant of rate. In addition, a chimera consisting of SP1/NC P4–P1 and MA/CA P1'–P4' residues (ATIM ↓ PIVQ) abolished processing by wild type and MDR proteases, and the reciprocal chimera consisting of MA/CA P4–P1 and SP1/NC P1'–4' (SQNY ↓ IQKG) was cleaved only by one of the MDR proteases. These results suggest that complex substrate interactions both beyond the active site of the enzyme and across the scissile bond contribute to defining the rate of processing by the HIV-1 PR.

Proteolytic cleavage of the HIV-1 polyproteins Gag and Gag-Pro-Pol by the viral protease (PR)² is an essential step to produce an infectious virus particle. Proteolysis occurs concomitantly with budding of the virus particle, and this processing releases mature virion proteins (matrix (MA), capsid (CA), spacer peptide 1 (SP1), nucleocapsid (NC), spacer peptide 2 (SP2), and p6) from Gag, and in addition the viral enzymes, protease, RT, and integrase (IN) from Gag-Pro-Pol (1). This event is highly regulated by a number of elements such as the substrate sequence/conformation, the presence of spacer peptide domains, the interaction with RNA, and the context of the cleavage sites (2–7). However, the precise mechanisms governing the rates of the cleavage events are not fully understood.

The processing sites in the Gag polyprotein are cleaved at different rates, resulting in sequential Gag processing during the maturation process (3, 6, 8–10). In an experimental setting where the Gag polyprotein was synthesized *in vitro* and cleavage was initiated with the addition of purified protease, the initial cleavage occurs at the SP1/NC site and the cleavages at the MA/CA and SP2/p6 sites follows at an ~10-fold slower rate. The slowest cleavages then occur at the CA/SP1 and NC/SP2 (3). The cleavage at the CA/SP1 site is the slowest, with a cleavage rate of about 400-fold slower than that of the initial cleavage (3). A similar system using Gag-Pro-Pol has shown that the cleavage at the RT/IN site occurs at a rate similar to the cleavage at the MA/CA site, and cleavages at the two sites flanking the PR (TF/PR and PR/RT) occur at the final stage of the Gag-Pro-Pol processing (9). The cleavage event at each site takes place independently of the other sites when protease is provided in *trans* with the exception of the cleavage at the CA/SP1 site, which showed enhanced cleavage when the initial cleavage event at the downstream SP1/NC site was blocked by a mutation, M377I (3). Thus, the processing at the CA/SP1 site is neg-

* This work was supported, in whole or in part, by National Institutes of Health Grant P01-GM066524.

¹ To whom correspondence should be addressed: CB7295 Lineberger Bldg., Rm. 22-006, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599. Tel.: 919-966-5710; Fax: 919-966-8212; E-mail: risunc@med.unc.edu.

² The abbreviations used are: PR, protease; MA, matrix; CA, capsid; SP, spacer peptide; NC, nucleocapsid; MDR, multidrug-resistant; IN, integrase.

actively regulated by the presence of the SP1 domain after release from the NC domain, which may create an unfavorable environment for cleavage by the HIV-1 protease.

The viral protease is a homodimer with the active site at the dimer interface (11–13), whereas the cleavage sites in the Gag and Gag-Pro-Pol polyproteins are structurally asymmetric, sharing little amino acid sequence homology (14). Crystallographic studies using peptides corresponding to the cleavage sites within the Gag and Gag-Pro-Pol polyproteins have revealed that the HIV-1 protease achieves its substrate specificity in part by recognizing a conserved shape rather than a particular amino acid sequence (15). In these studies, all of the peptide substrates were shown to adopt an asymmetric extended β -strand conformation when bound in the active site of the enzyme creating a consensus volume termed the “substrate envelope.” Recently, Ozen *et al.* (16) demonstrated how a particular substrate that fits within the substrate envelope is influenced by both substrate dynamics and size. In this study, the volume of the CA/SP1, NC/SP2, and SP2/p6 substrates protruded beyond the substrate envelope more than expected based on their size. These substrates were highly dynamic, resulting in large deviations from the crystal structure and a worse fit within the substrate envelope. Consistent with their slow rate of cleavage, the CA/SP1 and NC/SP2 substrates were shown to score as being the most dynamic among the substrates analyzed (16).

The order of cleavage obtained from the kinetic studies using peptide substrates corresponding to the cleavage sites in Gag and Gag-Pro-Pol differs from the order of cleavage in the context of the full-length Gag polyprotein (3, 9, 17–22). Furthermore, the peptide substrates used in different studies did not reproduce the same order of cleavage (19, 21, 22), suggesting that additional determinants beyond amino acid sequence and local secondary structure of the cleavage sites are involved in Gag and Gag-Pro-Pol processing. One clear example is that in the context of full-length Gag polyprotein, the cleavage of the CA/SP1 is negatively affected by the initial cleavage at the SP1/NC site (3). There are also reported differences in both the order and the cleavage rate of peptide substrates when different lengths of peptides were used, implying that amino acids outside the active site groove may play a role in Gag and Gag-Pro-Pol cleavage. One example of the importance of the context surrounding cleavage sites in HIV-1 Gag processing was demonstrated in the study by Tritch *et al.* (5) using a protein substrate derived from the Gag polyprotein. When the P4–P4′ residues of the MA/CA site were replaced with the P4–P4′ residues of the SP1/NC site, the rate of cleavage was enhanced compared with that of the MA/CA site; however, the extent of enhancement was shown to be far less than that expected compared with cleavage of this site in its normal setting, suggesting that the context outside the local cleavage site is also involved in the efficient cleavage of the SP1/NC site. It has also been reported that processing site context, specifically sequences flanking the C terminus of the protease, can affect enzyme activity (23).

Although the context outside of the cleavage site has been suggested to play a role as a determinant for the substrate specificity of the HIV-1 protease, studies regarding contextual influ-

ences on HIV-1 processing have not been addressed extensively using folded protein substrates. Therefore, to examine the effect of context on HIV-1 processing, we conducted a comparative analysis of the Gag and Gag-Pro-Pol cleavage sites in a heterologous environment using a sensitive HIV-1 protease assay. In this assay, the MA/CA fusion protein modified with several alterations was used as a substrate. A tetracysteine motif (CCGPCC) was introduced within the N-terminal domain of CA to allow binding of the FAsH (fluorescein arsenical hairpin) reagent (24–26). In addition, to differentiate the size of the cleavage products from two pooled substrates, a GST tag was fused to the N terminus of MA, and the C-terminal domain of CA was truncated. By utilizing this substrate labeled with FAsH reagent, we were able to measure specific proteolysis in gel-based assays as a function of HIV-1 protease cleavage. When the residues P4–P4′ of the cleavage sites within Gag and Gag-Pro-Pol were placed at the MA/CA site, we observed changes in the order of the cleavage rate, indicating that the context outside the P4–P4′ residues plays a crucial role for efficient Gag and Gag-Pro-Pol cleavage. Changes in the order of the cleavage rates were also seen with a multi-drug resistant (MDR) protease. Combining the P4–P1 residues from the SP1/NC site and the P1′–P4′ residues from the MA/CA site resulted in a chimeric substrate that was not significantly cleaved by any of the proteases tested in this study, whereas the reciprocal construct was selectively cleaved by one of the MDR proteases but not by the wild type protease. These results indicate that more complex interactions than just those of the substrate amino acids within the active site of the protease are required for efficient Gag and Gag-Pro-Pol cleavage by the HIV-1 protease and that this new series of folded protein substrates offers a tool to examine protease specificity.

EXPERIMENTAL PROCEDURES

Constructs—The primers to construct plasmids used in this study were designed so that a His tag was introduced at the N terminus of each protein. The plasmid pARKz1k1–5LTR_{gag}, containing a fragment of 5′ LTR and the *gag* region of pNL-CH, an infectious molecular clone derived from the pNL4-3 clone of HIV-1 (27), was used as a template to amplify the full-length MA/CA coding region by PCR. The PCR product was digested with NdeI, which was introduced in the PCR primers, and cloned into the NdeI site of pET30b (Novagen, Madison, WI) to generate the pMA/CA precursor plasmid. A tetracysteine motif (CCGPCC) was introduced within the N-terminal domain of CA (His-87–Ala-92) by site-directed mutagenesis following the QuikChange method (Stratagene, La Jolla, CA) to create the final pMA/CA plasmid (see Fig. 1A). For pMA/CA Δ , two overlapping PCR fragments, one containing the coding region for a GST tag amplified from pET41b (Novagen) and the other containing the full-length MA/CA coding region amplified from pMA/CA, were used in an overlapping extension PCR. The resulting PCR product was cloned into the NdeI site of pET30b, and then the C-terminal domain of CA was truncated by substituting Ser-278 with a stop codon. The pMA/CA Δ -noTC was created in the same way that the pMA/CA Δ was generated except that the full-length MA/CA coding region was amplified from pARKz1k1–5LTR_{gag} to exclude the tetracysteine motif

from the CA region. To generate pMA/CA Δ -Y132I, a Y132I mutation was introduced at the P1 position of the MA/CA cleavage site by site-directed mutagenesis using the pMA/CA Δ as the template. For the constructs shown in Fig. 1B, eight codons representing the P4–P4' positions of the MA/CA cleavage site in the pMA/CA Δ were replaced with either the equivalent codons from the CA/SP1, SP1/NC, NC/SP2, SP2/p6, TF/PR, PR/RT, RT/IN, or chimeric cleavage sites using overlapping extension PCR. The SP1/NC P4–1 was generated by overlapping extension PCR so that the P4–P1 residues are derived from the SP1/NC cleavage site and the P1'–P4' residues are derived from the MA/CA cleavage site. The SP1/NC P1'–4' was generated in the same way that the SP1/NC P4–1 was generated except that the P4–P1 residues are derived from the MA/CA cleavage site and the P1'–P4' residues are derived from the SP1/NC cleavage site. The structure of each of the constructs was confirmed by DNA sequence analysis.

Expression and Purification of HIV-1 PR Substrates—Expression of recombinant proteins in *Escherichia coli* BL21 (DE3) was carried out by a modified version of established protocols (28, 29). Briefly, recombinant proteins were expressed in MagicMedia (Invitrogen) for 7 h, and the cells were collected by centrifugation. The cell pellet was lysed in Tris-buffered saline (25 mM Tris base, 3 mM KCl, and 140 mM NaCl) at pH 7.5 with 1 mM DTT and 1% Triton X-100 and then sonicated. Following clarification by centrifugation, the recombinant proteins containing a His₆ tag at the N termini were purified from the soluble fraction by affinity chromatography using a nickel-chelating column (Novagen). The purified recombinant proteins were dialyzed against 20 mM sodium acetate (pH 7.0), 1 mM EDTA, 2 mM DTT, and 10% glycerol (30). Protein concentrations were determined by using the Bradford assay (Bio-Rad, Hercules, CA), and the purity of the purified proteins was analyzed by protein staining using SimplyBlue SafeStain (Invitrogen) after SDS-PAGE (31).

Protease Expression and Purification—The HIV-1 protease was expressed in *E. coli* TAP 106 cells and purified from inclusion bodies as previously described (32). Briefly, the inclusion body centrifugation pellet was dissolved in 50% acetic acid followed by another round of centrifugation to remove impurities. Size exclusion chromatography was used to separate high molecular weight proteins from the protease. The protein was refolded in 50 mM sodium acetate at pH 5.5, 5% ethylene glycol, 10% glycol, and 5 mM DTT.

FLaSH Labeling Reactions—The FLaSH reagent (Invitrogen) was incubated overnight (~15 h) with the MA/CA-protein substrate in a 1:1 molar ratio at room temperature (33) in proteolysis buffer (50 mM sodium acetate, pH 7.0, 150 mM NaCl, 1 mM EDTA, 2 mM 2-mercaptoethanol, and 10% glycerol). A total concentration of 6 μ M of the protein substrates was used for FLaSH labeling in a total volume of 228 μ L.

Gel-based Protease Assay by HIV-1 PR—To compare processing efficiency of the heterologous substrates (cleavage sites introduced within the MA/CA site) with the MA/CA substrate, 3 μ M of each of two substrates (a heterologous substrate plus the MA/CA substrate) was labeled with 6 μ M FLaSH reagent in a tube with a total volume of 228 μ L. An aliquot of 24 μ L was taken for the 0-min time point before the addition of 0.1 μ M

HIV-1 PR to the substrate. Proteolysis was performed at 30 °C, and aliquots of 24 μ L were taken at various time points, and the reaction was stopped by adding 4 \times SDS-PAGE loading buffer. After the last time point, the protein in the aliquots was resolved by SDS-PAGE without a boiling step. The concentration of a reducing reagent, 2-mercaptoethanol, in the SDS-PAGE loading buffer was kept at 2 mM to keep the FLaSH reagent coupled to the substrate. The gels were briefly rinsed with water, and the fluorescently labeled protein bands were visualized by fluorescence imaging using a Typhoon 9400 (GE Healthcare/Amersham Biosciences) with excitation at 488 nm and emission at 526 nm. The relative quantitation of the visualized protein bands was performed by using the image analysis software ImageQuant TL (GE healthcare). To calculate the relative rate of individual cleavage site, the percentage amount of the cleavage products within the linear range was plotted as a function of the time. Briefly, the time points representing either less than or equal to 20% cleavage were used for linear regression, and then the slope value of the individual cleavage site was used to determine cleavage rate relative to the MA/CA site. Unless specified, all of the labeling and proteolysis reactions were performed at pH 7.0 because of the dependence of the FLaSH reagent on pH. For binding specificity of the FLaSH reagent to a tetracysteine motif, the labeling reaction of the 3 μ M substrate (MA/CA Δ or MA/CA Δ -noTC) was performed in the absence or presence of 3 μ M FLaSH reagent at pH 7.0. Proteolysis was performed in the absence or presence of the 2 μ M HIV-1 PR at 30 °C for 3 h, and the resulting protein products were subjected to SDS-PAGE. The fluorescent protein bands were visualized by Typhoon 9400, and the same gel was stained with SimplyBlue SafeStain for Coomassie staining. For the effect of pH on HIV-1 PR activity, 3 μ M substrate (MA/CA Δ) was used for labeling reaction at pH 5.5, 6.0, 6.5, or 7.0 following the protocol described above.

Estimation of van der Waals Potential and Hydrogen Bonding—For the structural analyses, crystal structures were used when available. When an experimentally determined structure was not available, the structure of the substrate complex that was most similar in sequence to the target substrate was used as the template. To this template, the mutations were introduced *in silico* using Maestro from Schrodinger Suite 2011 (Schrodinger, Portland, OR). All of the structures (experimental and modeled) were prepared for the structural analyses using the Protein Preparation Wizard of Maestro. In the preparation step, hydrogen atoms were added, and the missing side chains were filled in and refined. The hydrogen bonding network was optimized at neutral pH using the exhaustive sampling option. Orientations of the crystallographic waters were also sampled for the optimal hydrogen bonding network. Energy minimization was performed on the coordinates of the hydrogen atoms only, followed by a second restrained minimization on the whole system until root mean square deviation converged to 0.3 Å for crystal structures and 0.5 Å for the modeled structures. After preparation, each structure was analyzed for the protease-substrate hydrogen bonds and van der Waals contacts. The total number of hydrogen bonds was counted by using the default geometric criteria in Maestro (2.5 Å maximum distance, 120° minimum donor angle, and 90° minimum acceptor angle). Protease-substrate van der Waals interactions were calculated by a

Effect of Context on HIV-1 Gag Processing

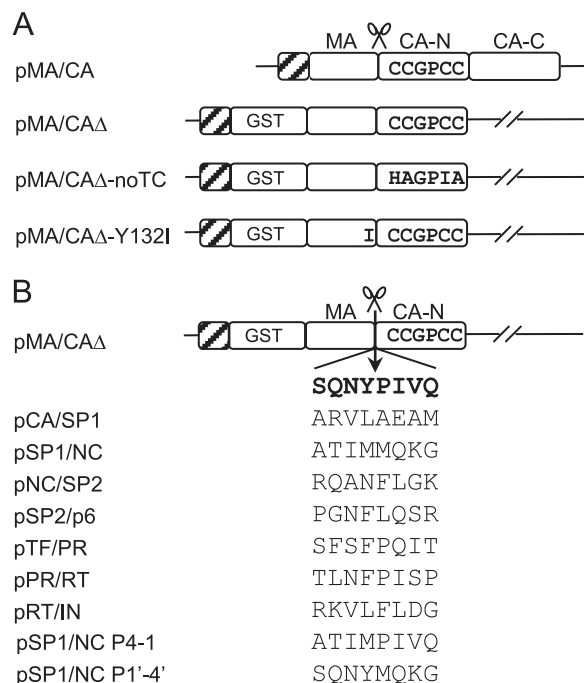


FIGURE 1. **Schematic diagram of MA/CA constructs.** *A*, the structure of the constructs in backbone of the *E. coli* expression vector pET30b. The parental construct, MA/CA, contains an intact MA and CA domain. The constructs MA/CA Δ , MA/CA Δ -noTC, and MA/CA Δ -Y132I include alterations with the addition of GST tag at the N terminus of the MA domain and the truncation of C-terminal domain of CA. The MA/CA Δ -Y132I has an Ile substitution at the P1 position of the MA/CA cleavage site that blocks cleavage at the site. The tetracysteine motif (CCGPCC) is introduced within the N-terminal domain of CA, and the *hatched boxes* at the N terminus represent the His₆ tag. The cleavage site between MA and CA is indicated. *B*, Gag and Gag-Pro-Pol cleavage sites in the context of the MA/CA. The amino acid sequences of the cleavage sites within the Gag and Gag-Pro-Pol region substituting the P4–P4' residues of the MA/CA site (SQNY \downarrow PIVQ) in the MA/CA Δ construct are shown. The chimeric SP1/NC P4–1 contains P4–P1 derived from the SP1/NC site and P1'–P4' derived from the MA/CA site, whereas the SP1/NC P1'–4' contains P4–P1 derived from the MA/CA site and P1'–P4' derived from the SP1/NC site.

simplified Lennard-Jones potential function as described in detail elsewhere (16).

RESULTS

Design of Substrates for Gel-based HIV-1 Protease Assay Employing FAsH Reagent—To study the effect of context/environment on the processing of the HIV-1 Gag and Gag-Pro-Pol cleavage sites by the HIV-1 PR, we developed a protease assay in which a Gag protein substrate is labeled with a FAsH reagent, a nonfluorescent biarsenical derivative of fluorescein that becomes fluorescent upon binding to its target (26). In this assay, a folded MA/CA protein containing an intact MA/CA cleavage site was used as a prototype substrate with the following modifications. A tetracysteine motif (CCGPCC) was introduced into a surface loop region within the N-terminal domain of CA (amino acids 87–92), near the position previously used to label intact virus (34), to allow for FAsH reagent binding, resulting in a substrate designated as MA/CA (Fig. 1A). The substrate MA/CA Δ (GST-MA-CA Δ CTD) was made by fusing a GST tag to the N terminus of MA to make the substrate larger and truncating the C-terminal domain (CTD) of CA to decrease the size of the cleaved labeled product (Fig. 1A). The MA/CA Δ -noTC is the same substrate as the MA/CA Δ except that the

MA/CA Δ -noTC lacks the tetracysteine motif needed for FLAsH reagent binding. To introduce a heterologous cleavage site into the context of the MA/CA cleavage site, the MA/CA cleavage site within MA/CA Δ was replaced with each of the seven cleavage sites from the Gag and Gag-Pro-Pol polyprotein, individually generating substrates with the cleavage sites CA/SP1, SP1/NC, NC/SP2, SP2/p6, TF/PR, PR/RT, and RT/IN, as shown in Fig. 1B. Because of the differences in size of the two substrates, MA/CA and MA/CA Δ , and the labeled cleavage products, CA and CA Δ , this strategy enables us to perform proteolysis reactions in the presence of the MA/CA substrate as an internal control and to resolve the cleavage products by SDS-PAGE.

Efficient Processing of MA/CA Δ by HIV-1 PR and Specific Binding of FAsH Reagent to Tetracycline Motif—Because the MA/CA fusion protein was modified for the assay, we first compared the efficiency of cleavage of the altered substrate MA/CA Δ (≈ 56 kDa) with that of the MA/CA protein (≈ 40 kDa) in the gel-based assay, placing equal molar amounts of each substrate into the reaction. Because the two substrates and their cleavage products differ in size and hence can be resolved by SDS-PAGE, we were able to mix the two substrates and carry out the protease assay together. Fig. 2A shows the cleavage products over a 3-h time period visualized by fluorescence imaging. The N-terminal domain of CA harbors the tetracycline motif where the FAsH reagent binds; thus only the proteins containing the N-terminal domain of CA are detected by fluorescence imaging. Both substrates were processed by the HIV-1 protease at a similar rate over time as assessed by the band intensity of the cleaved CA products, the full-length CA released from MA/CA, and the CA Δ (≈ 15 kDa) released from MA/CA Δ (Fig. 2, A and B). Thus, the alterations made to the MA/CA protein did not affect its cleavage efficiency by the protease nor result in significant artifactual degradation of the protein or cleavage at fortuitous sites. We also tested cleavage efficiency of the SP2/p6 site (a site that is cleaved slowly, see below) in both substrates, MA/CA and MA/CA Δ , and we observed very similar cleavage rates with each substrate (data not shown), confirming that the alterations we made to MA/CA did not affect cleavage at the MA/CA site. The faint band running slightly slower than CA protein, noted with an *arrow* in Fig. 2A, seems to be an unknown contaminant specifically of the MA/CA Δ prep that is labeled with the FAsH reagent; its nature is unclear because it is present in the uncleaved prep, and the band intensity does not change in the presence of protease.

Studies done *in vivo* and *in vitro* have previously demonstrated that the FLAsH reagent binds a tetracysteine motif with high affinity and specificity (26), and proteins labeled with the FLAsH reagent have been analyzed by SDS-PAGE (24). However, in this previous report the detection of such protein bands in a gel was not optimal, which may have been due to the use of crude *E. coli* lysates for FLAsH labeling. To test the specificity of the FLAsH reagent binding to the tetracysteine motif, two substrates, one harboring a tetracysteine motif (MA/CA Δ) and the other lacking a tetracysteine motif (MA/CA Δ -noTC), were visualized by both Coomassie staining and fluorescence imaging in the absence or presence of FLAsH reagent and protease cleavage (Fig. 2C). In the presence of protease, Coomassie Blue staining showed that both MA/CA Δ and MA/CA Δ -noTC gen-

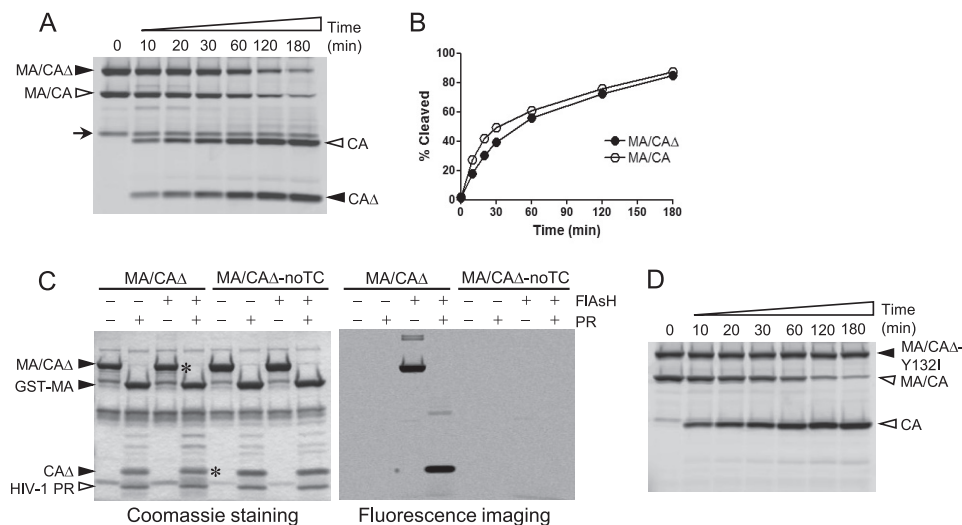


FIGURE 2. Efficient processing of the MA/CA Δ by HIV-1 PR and specific binding of the FIAsh reagent to the tetracycysteine motif. *A*, both substrates MA/CA Δ and MA/CA were labeled with the FIAsh reagent in the same reaction, and proteolysis of the mixed substrates was performed at 30 °C by adding 0.1 μ M HIV-1 PR. Equal molar amounts of the two substrates were used, and the same molar ratio between the FIAsh reagent and total amounts of the protein were used. Aliquots were taken at the various time points shown above each lane and resolved by SDS-PAGE. The aliquot for the 0-min time point was taken before the enzyme was added. Fluorescently labeled proteins were visualized by fluorescence imaging. *Filled arrowheads* indicate MA/CA Δ substrate and its cleavage product (CA Δ), and *open arrowheads* indicate MA/CA substrate and its cleavage product (CA). *B*, the band intensity of the cleavage products, CA Δ and CA, in *A* was quantified and plotted as a function of the time. *C*, the substrate MA/CA Δ carrying a tetracycysteine motif and the MA/CA Δ -noTC lacking a tetracycysteine motif were subjected to labeling reactions in the absence or presence of FIAsh reagent. After the labeling reaction, the proteolysis reaction was performed in the absence or presence of 2 μ M HIV-1 PR. The cleavage pattern of the substrates is visualized by Coomassie Blue staining (*left*) and fluorescence imaging (*right*). The bands indicated with an asterisk are visualized by fluorescence imaging. *D*, cleavage of the MA/CA Δ -Y132I is displayed over time to show specific processing of the MA/CA site by the HIV-1 PR.

erated the same two cleavage products, p40 GST-MA and the p15 truncated CA protein, CA Δ . The same gel visualized by fluorescence imaging revealed that only the bands (indicated by the asterisks) of the proteins harboring a tetracycysteine motif had the FIAsh reagent. The fact that the MA/CA Δ -noTC protein did not show any nonspecific binding of the FIAsh reagent to the protein confirms that the interaction between the tetracycysteine motif and the FIAsh reagent is highly specific. The band intensity of FIAsh-labeled protein visualized by fluorescence imaging depends on the number of the molecules rather than the size of the molecules, providing a broad and linear detection range. When we compared the sensitivity of detection between several methods, the gel-based assay using FIAsh-labeled protein showed sensitivity of detection comparable with that of silver staining (data not shown). Finally, to show specificity of protease cleavage, we combined the wild type substrate with a substrate where the P1 tyrosine was replaced with isoleucine to block protease cleavage. When these substrates were incubated together with protease, only the substrate with the wild type cleavage site was cleaved (Fig. 2D).

Lower pH (<6.0) Is Not Optimal for Protease Cleavage of Protein Substrate MA/CA Δ —The gel-based protease assay was performed at pH 7.0 because the FIAsh reagent becomes non-fluorescent at pH values below 6.0 (35). However, most enzymatic analyses of the HIV-1 protease using peptide substrates have been performed at lower pH, typically near pH 5.0, and the rates of cleavage of the MA/CA and CA/SP1 sites have been shown to be accelerated at pH 5.0 compared with pH 7.0 when peptide substrates were used (3). Although HIV-1 protease assays using full-length Gag polyprotein have been performed at pH 7.0 (3, 5), the cleavage rate of the CA/SP1 in the context of the full-length Gag polyprotein was shown to be significantly

increased at lower pH relative to cleavage of MA/CA (3). Therefore, to examine whether the processing of the MA/CA substrate by the HIV-1 protease at pH 7.0 is optimal compared with the cleavage at pH 5.5, the enzyme activity of the protease at different pH values, pH 5.5, 6.0, 6.5, and 7.0, was compared. Surprisingly, in the context of a large protein substrate, the protease cleaved the MA/CA site more efficiently at pH values higher than 6.0, although no significant difference was seen between pH 6.0 and 7.0 (Fig. 3, A and B). As expected, band intensity was uniformly decreased at the pH 5.5 because of the FIAsh reagent becoming nonfluorescent at lower pH (for both the precursor and the product). In contrast with previous studies using peptide substrates, the cleavage rate of the protein substrate by the protease was dramatically accelerated at higher pH (approximately a 10-fold increase at pH values above 6.0) compared with the rate of cleavage observed at pH 5.5 (Fig. 3C). Thus, optimal pH in this protease assay using MA/CA substrate is between 6.5 and 7.0, which is more consistent with physiological pH.

Impact of Heterologous Context on Cleavage Rates of Processing Sites within Gag and Gag-Pro-Pol Polyproteins—The processing sites in the Gag polyprotein are cleaved at different rates resulting in sequential Gag processing during the maturation process to produce an infectious virion. In a previous *in vitro* study (3, 9, 10), the processing rates of the different cleavage sites in the Gag precursor gave the following relative rates of cleavage (setting the MA/CA cleavage rate to 1): SP1/NC 14X, SP2/p6 1.6X, MA/CA 1X, and CA/SP1 and NC/SP2 0.04X. In the current study, we replaced the eight amino acids (P4–P4') at the MA/CA cleavage site within the MA/CA Δ substrate with the cleavage site sequences from the other Gag and Gag-Pro-Pol cleavage sites: CA/SP1, SP1/NC, NC/SP2, SP2/p6, TF/PR,

Effect of Context on HIV-1 Gag Processing

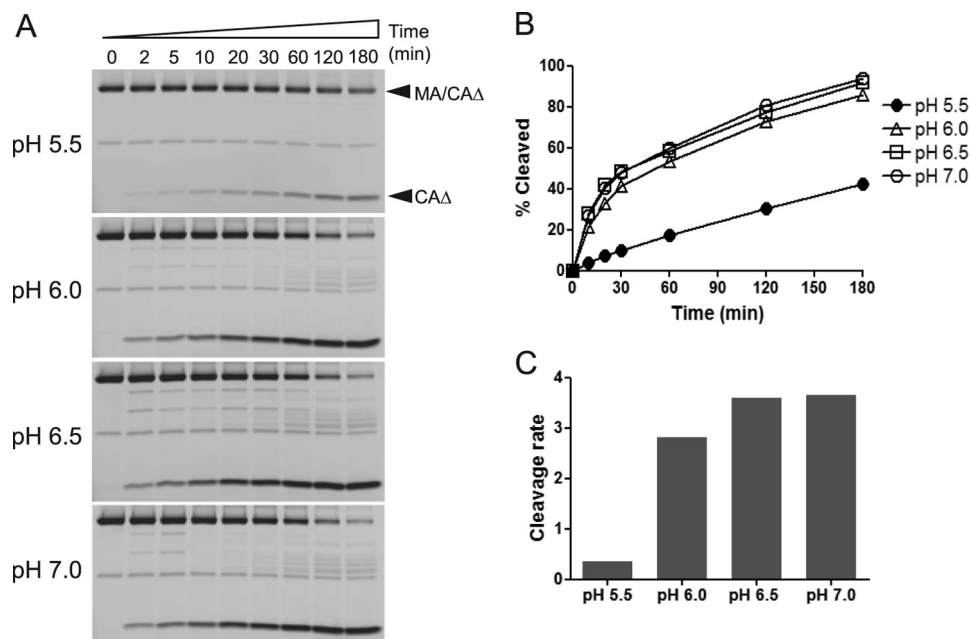


FIGURE 3. Effect of pH on proteolysis of MA/CAΔ by HIV-1 PR. A, the labeling and proteolysis reactions of the substrate MA/CAΔ were performed at pH 5.5, 6.0, 6.5, or 7.0 as described in Fig. 2A. Fluorescently labeled proteins were visualized by fluorescence imaging. Filled arrowheads indicate the substrate before cleavage (MA/CAΔ) and after cleavage (CAΔ). B, the band intensity of the cleavage products, CAΔ, shown in A was quantified and plotted as a function of the time. C, cleavage rates were calculated using the time points in which the percentage amounts of the cleavage products are within the linear range ($\leq 20\%$). The slope value obtained from linear regression analysis was used to compare the cleavage rate among different pH levels.

PR/RT, or RT/IN. Although no changes were made in amino acid sequences required for processing, the transposition of these sequences resulted in the cleavage sites being placed in a heterologous environment in the context of the Gag precursor. If the context outside of each cleavage site plays a role in determining cleavage rates, we would expect to see changes in the relative order of Gag cleavage. In this design, the initial rate of cleavage defines the specificity constant (k_{cat}/K_m), and the difference in the initial rate of the two different substrates, when present at equal concentrations, defines the difference in the specificity constant. In the following experiments the initial rate of cleavage of the wild type MA/CA substrate was set as 1.

As noted above, these experiments were designed so that the wild type protein (MA/CA) and the protein with the alternative cleavage site had different sizes as both substrate and cleavage product, which allowed us to mix the two substrates and carry out the cleavage reaction in a single tube. Thus, the wild type substrate, MA/CA (Fig. 1A), was included in each reaction as an internal control to measure the relative cleavage rate. Fluorescence images of the cleavage reactions for the CA/SP1, SP1/NC, and RT/IN and the resulting time course of the cleavage for all of the substrates are shown in Fig. 4 (A and B), respectively. When the cleavage rates were compared with that of the MA/CA site, we observed significant differences in the rate of cleavage for some of the different Gag and Gag-Pro-Pol cleavage sites (Fig. 4, B and C). As can be seen in Fig. 4C, the sites can be grouped into those cleaved within 3-fold of the rate of MA/CA and those that were cleaved much more slowly. MA/CA and RT/IN have similar cleavage rates in the context of Gag-Pro-Pol (9) and also similar rates in the context of MA/CA (Fig. 4C). SP1/NC has a 10-fold higher cleavage rate in its homologous position, but when transposed the rate was slower

than MA/CA, but cleavage was still detectable (Fig. 4C). CA/SP1 is cleaved slowly, but this rate is dictated by prior removal of the NC domain by cleavage at SP1/NC; when the SP1/NC cleavage site is blocked then CA/SP1 cleavage occurs at a rate that is similar to MA/CA (3), which was also the case when the site is placed at the MA/CA site (Fig. 4C). NC/SP2, TR/PR, and PR/RT are all cleaved slowly in the context of Gag and Gag-Pro-Pol (3, 9, 10) and were cleaved poorly when transposed into the MA/CA site (Fig. 4C). The most anomalous rate is that of SP2/p6, which is similar to MA/CA in the context of Gag (3) but was cleaved very poorly when the site was transposed (Fig. 4C). However, the magnitude of the decrease of the rate of cleavage of SP1/NC and SP2/p6 is similar compared with the rate of cleavage in Gag, suggesting that especially for these two sites their context in Gag plays a significant role in enhancing the rate of cleavage.

Inhibitor Resistance Mutations in HIV-1 Protease Can Change Processing Site Preferences—The presence of multiple mutations in the protease region can reduce binding affinity to protease inhibitors leading to resistance to the protease inhibitors. It is possible that these protease variants also alter substrate specificity (36, 37). To see whether MDR HIV-1 protease variants can affect the order of the rates of the Gag and Gag-Pro-Pol cleavage, we tested two MDR HIV-1 protease variants in our protease assay: PR_{G2} containing amino acid substitutions L10I, G48V, I54V, L63P, and V82A; and PR_{G4} containing amino acid substitutions L10I, L63P, A71V, G73S, I84V, and L90M. As shown in the Fig. 5, the PR_{G2} enzyme had relative rates of cleavage of the different substrates that were overall similar to the wild type enzyme with the exception of the SP1/NC site, which was cleaved at a six times faster rate than the MA/CA site and now more similar to its rate in full-length Gag (3), and the

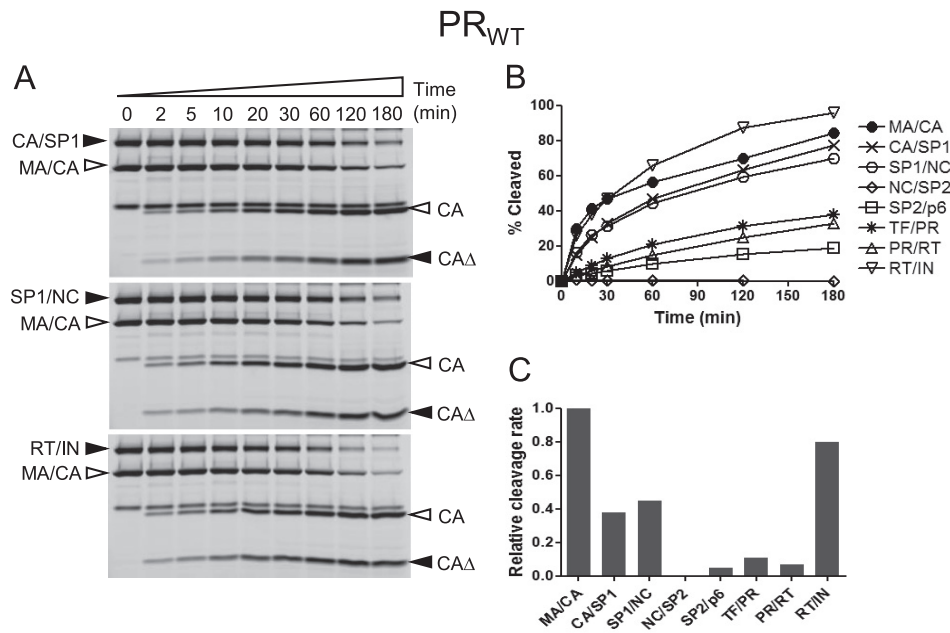


FIGURE 4. Processing of the Gag and Gag-Pro-Pol cleavage sites in the context of the MA/CA by PR_{WT}. A, fluorescence images displaying the processing of the CA/SP1, SP1/NC, and RT/IN sites in the heterologous MA/CA context are shown. The labeling and proteolysis reactions of the substrates were performed at pH 7.0 as described in Fig. 2A. The MA/CA substrate was included in all of the cleavage reaction as an internal control, which allows direct comparison of cleavage rates between the reactions. Filled arrowheads indicate MA/CA Δ substrates containing the heterologous cleavage site and its cleavage product (CA Δ), and open arrowheads indicate MA/CA substrate and its cleavage product (CA). B, the percentage amounts of the cleavage products are shown over time. The standard deviation of the MA/CA site is shown. C, cleavage rates of the processing sites within the Gag and Gag-Pro-Pol region are shown, relative to that of the MA/CA site, which was set at 1.

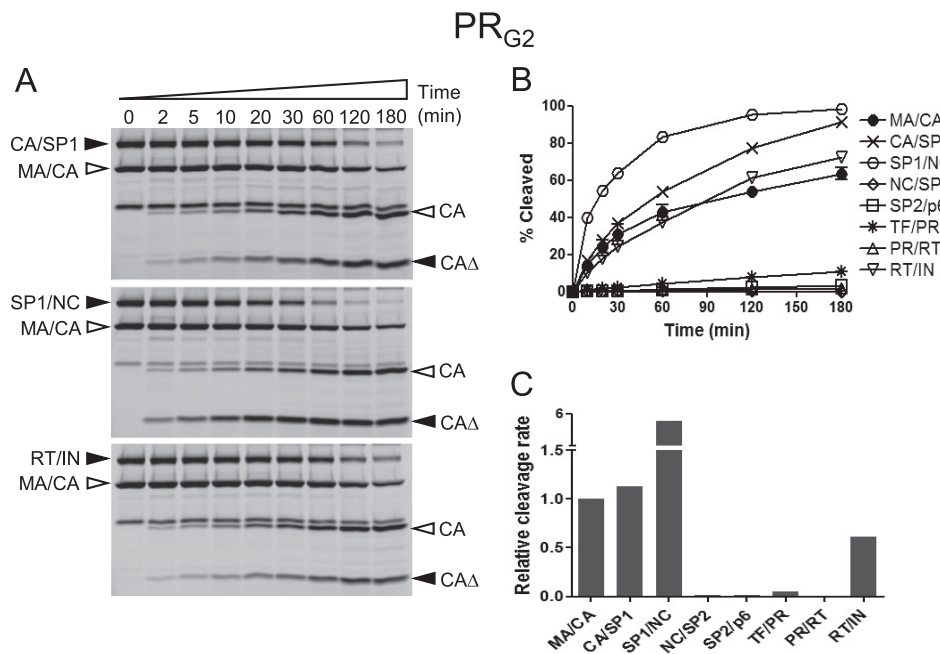


FIGURE 5. Processing of the Gag and Gag-Pro-Pol cleavage sites in the context of the MA/CA by PR_{G2}. A, fluorescence images depicting the processing of the CA/SP1, SP1/NC, and RT/IN sites in the heterologous MA/CA context are shown. The labeling and proteolysis reactions of the substrates were performed at pH 7.0 as described in Fig. 2A. Filled arrowheads indicate MA/CA Δ substrates containing the heterologous cleavage site and its cleavage product (CA Δ), and open arrowheads indicate MA/CA substrate and its cleavage product (CA). B, the percentage amounts of the cleavage products are shown over time. The standard deviation of the MA/CA site is shown. C, cleavage rates of the processing sites within the Gag and Gag-Pro-Pol region are shown, relative to that of the MA/CA site, which was set at 1.

CA/SP1 site, which was cleaved moderately faster compared with the MA/CA site. In contrast, the PR_{G4} variant resulted in relative rates that were similar to those of the PR_{WT} with the exception of the RT/IN site, which was modestly reduced (Fig. 6). Thus, the two MDR proteases showed distinct but mostly modest effects on substrate processing except for one of them

(PR_{G2}) changing the relative rate of cleavage of one site (SP1/NC) by 10-fold.

Chimeric Cleavage Site Composed of P4–P1 Amino Acids from SP1/NC Site and P1'–P4' Amino Acids from MA/CA Site Selectively Abolishes Processing by HIV-1 Proteases—Both MA/CA and SP1/NC sites are efficiently cleaved either in the

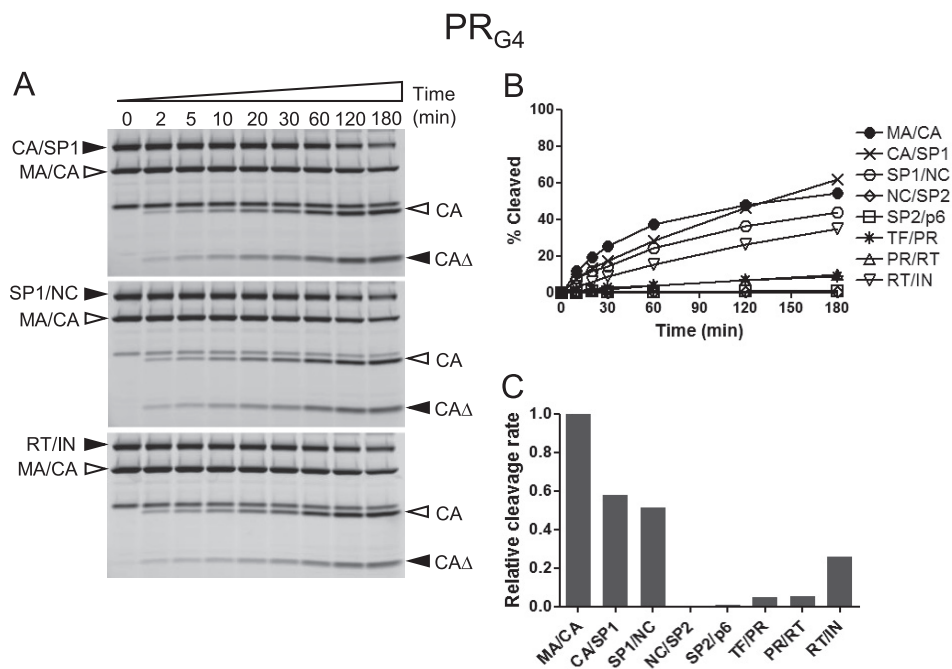


FIGURE 6. Processing of the Gag and Gag-Pro-Pol cleavage sites in the context of the MA/CA by PR_{G4}. A, fluorescence images displaying the processing of the CA/SP1, SP1/NC, and RT/IN sites in the heterologous MA/CA context are shown. The labeling and proteolysis reactions of the substrates were performed at pH 7.0 as described in Fig. 2A. Filled arrowheads indicate MA/CA Δ substrates containing the heterologous cleavage site and its cleavage product (CA Δ), and open arrowheads indicate MA/CA substrate and its cleavage product (CA). B, the percentage amounts of the cleavage products are shown over time. The standard deviation of the MA/CA site is shown. C, cleavage rates of the processing sites within the Gag and Gag-Pro-Pol region are shown, relative to that of the MA/CA site, which was set at 1.

context of the Gag polyprotein or in the context of the peptide substrates, and these two sites were differentially recognized by the wild type and two MDR proteases (Figs. 4–6). Thus, to explore substrate specificity further, we investigated two chimeric substrates: SP1/NC P4–1 and SP1/NC P1'–4'. SP1/NC P4–1 harbors P4–P1 derived from the SP1/NC site and P1'–P4' derived from the MA/CA site, whereas SP1/NC P1'–4' contains P4–P1 derived from the MA/CA site and P1'–P4' derived from the SP1/NC site. When these two chimeric substrates were tested for cleavage efficiency by the three different proteases, PR_{WT}, PR_{G2}, and PR_{G4}, the chimeric substrates were poorly cleaved by all of the enzymes (Fig. 7, A and B). The one exception was the SP1/NC P1'–4' chimera, which was cleaved faster than the MA/CA site but slower than the SP1/NC site, when cleaved by the PR_{G2} enzyme. Although the P1' to P4' amino acids accounted for much of the enhanced efficiency of the PR_{G2} protease for the SP1/NC site, in most cases the chimeras were poor substrates compared with their parental sequences, implying that the amino acids on either side of the scissile bond are interdependent for optimal processing.

Substrate specificity of the HIV-1 protease is determined by a number of elements responsible for the molecular interactions between protease and substrates. Thus, it is possible that changes in the molecular interactions between protease and the substrate caused by the alterations made in the SP1/NC P4–1 create a substrate that cannot be cleaved by the HIV-1 protease. To explore this possibility, the total number of hydrogen bonds and the van der Waals contacts between the protease and the chimeric substrates (P4–P4') within the structure of the protease-substrate complexes were analyzed using previously determined structures (16) and models of the chimeric sub-

strates derived from these structures. The van der Waals interactions reflect the favorable protease-substrate contacts caused by local packing in the hydrophobic binding groove, whereas hydrogen bonds play a crucial role in substrate specificity (15, 16). When we compared the results from the chimeric substrates to those from the parental substrate sites MA/CA and SP1/NC, the SP1/NC P4–1 had fewer total hydrogen bonds (Fig. 7C) and less van der Waals contact potential than the parental substrates (Fig. 7D). The SP1/NC P1'–4' substrate exhibited relatively high scores in these measures of molecular interactions between the protease and the substrate compared with the other substrates (Fig. 7, C and D); however, it was cleaved poorly by the PR_{WT} and PR_{G4} proteases (Fig. 7, A and B). The structural data seem to reflect these cleavage patterns only when both the substrate dynamics within the binding site and the impact of the residues surrounding the processing sites on the protease-substrate interactions are incorporated into the molecular modeling studies.

DISCUSSION

The HIV-1 PR has been an important target for a successful group of HIV-1 inhibitors because of its essential role in the proteolytic processing of HIV-1 Gag and Gag-Pro-Pol polyprotein for the formation of infectious virus. However, much still remains to be learned about the determinants of HIV-1 PR specificity in the context of the processing/assembly cascade. Most of our knowledge regarding PR specificity has come from the studies using peptides as substrates representing the Gag and Gag-Pro-Pol cleavage sites, which cannot account for possible elements residing outside of the local cleavage site sequences, P4–P4'. There has been a disagreement between

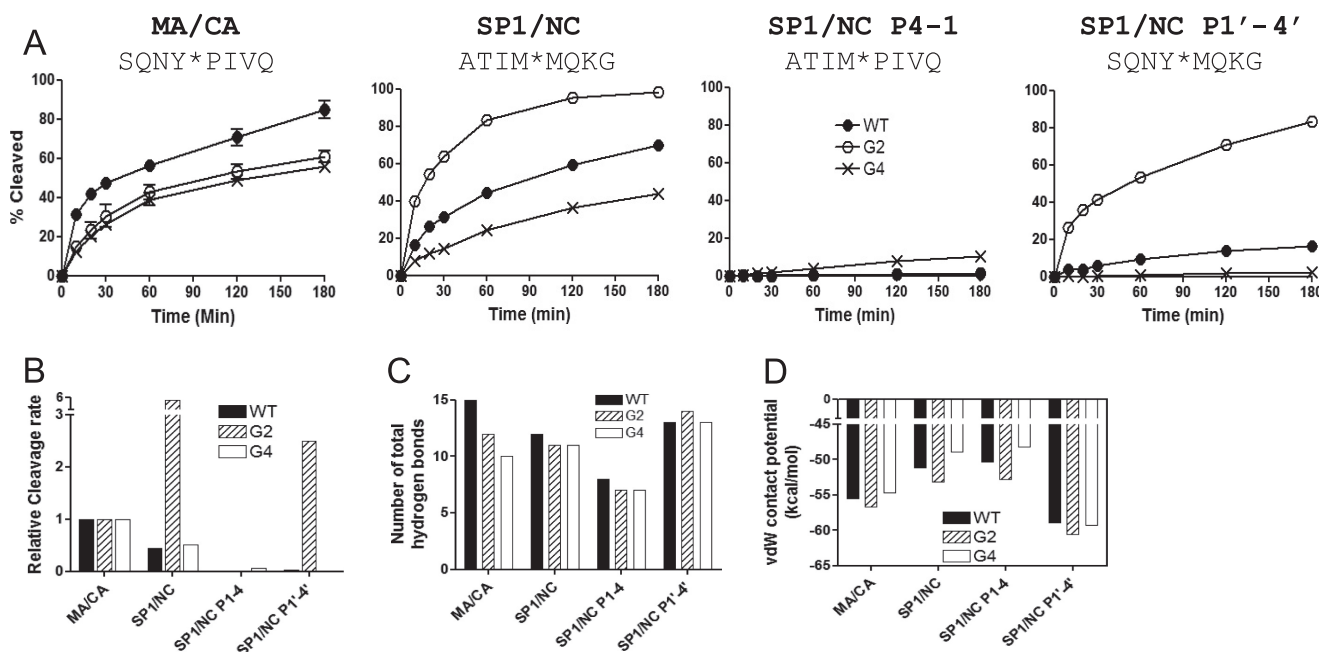


FIGURE 7. Reduced processing of the chimeric cleavage sites composed of the amino acids from the SP1/NC and MA/CA sites. *A*, the percentage amount of the cleavage product of the MA/CA, SP1/NC, SP1/NC P4-1, and SP1/NC P1'-4' sites processed by PR_{WT}, PR_{G2}, or PR_{G4} are shown as a function of time. The SP1/NC P4-1 contains P4-P1 residues derived from the SP1/NC site and P1'-P4' residues derived from the MA/CA site. The SP1/NC P1'-4' contains P4-P1 residues derived from the MA/CA site and P1'-P4' residues derived from the SP1/NC site. The amino acid sequence of the cleavage sites (MA/CA, SP1/NC, SP1/NC P4-1, and SP1/NC P1'-4') is shown. *B*, cleavage rates of the MA/CA, SP1/NC, SP1/NC P4-1, and SP1/NC P1'-4' sites relative to that of the MA/CA site are shown. *C*, number of total hydrogen bonds between the protease and substrates. *D*, calculated van der Waals contact potentials between the protease and substrates.

the order of cleavage obtained from the studies with peptide substrates and the order of cleavage in the context of full-length Gag polypeptide. Moreover, the order of cleavage of various cleavage sites ranked by different studies with peptide substrates is not consistent (19, 21, 22). One discrepancy among these studies is the length of the peptide used, which suggests that the local environment surrounding the heptapeptide cleavage site may play a role in determining substrate specificity of the HIV-1 PR, although methodological differences may also account for the reported rate differences. Thus, in an attempt to explore the question of context further, we performed a comparative analysis of the effect of context on HIV-1 processing by introducing Gag and Gag-Pro-Pol cleavage sites into the heterologous context of the Gag MA/CA cleavage site. By labeling the protein substrate, MA/CAD, with a FLA_{SH} reagent that binds the tetracysteine motif introduced within the NTD region of CA, we were able to detect and measure proteolysis of the substrate by the HIV-1 PR. In addition, we were able to manipulate the length of both the upstream and downstream segments of the products to allow two substrates and their labeled products to be resolved together by SDS-PAGE, which enabled us to run the wild type sequence as an internal control/standard in each reaction. Replacing the MA/CA cleavage site with cleavage sites from the Gag and Gag-Pro-Pol polypeptides resulted in similarities and differences in the expected rates of cleavage, revealing that the context surrounding at least three of the cleavage sites plays a major role in determining the rate of cleavage by the HIV-1 PR. In addition, the use of MDR HIV-1 PR variants, PR_{G2} and PR_{G4}, also selectively affected the rates of cleavage. The use of a fluorescent tag has also allowed us to develop a protease assay based on cleavage of a folded protein

using fluorescence anisotropy, which is amenable to large scale screens directed at inhibitors of either the protease or the substrate.³

Our goal in using a folded protein substrate was to be able to examine the cleavage site requirements in the context of the larger protein. We found that with the large MA/CA protein substrate, the optimal pH for cleavage by the protease was between 6.5 and 7.0. This result differs from earlier studies using peptides as substrates, which showed an optimum pH below 6 for proteolysis (3, 17, 38–40). In an *in vitro* assay using radiolabeled Gag polypeptide as a substrate, Pettit *et al.* (3) observed that the CA/SP1 site was cleaved about 20-fold faster at pH 5 compared with a reaction run at pH 7, measured as the rate relative to the MA/CA site. For synthetic peptide substrates, the K_m values were higher at pH 7 relative to lower pH, suggesting that peptides interact less well with the HIV-1 protease at pH 7 than at lower pH (38). The higher rate of cleavage at neutral pH that we observed is more consistent with the expectation that the pH in the virion should be the same as that in the cell.

Our previous work looking at the relative rates of the cleavage in the Gag and Gag-Pro-Pol polypeptide (3, 9) provide a context for interpreting the rates of cleavage of the same sites when placed in the MA/CA site. We propose that three phenomena are at work in defining the rates of cleavage when the sites are placed in this heterologous site. First, the rate of cleavage for one set of sites is largely determined by the sequence of the cleavage site such that the rate of cleavage is similar at the

³ S.-K. Lee, C. A. Schiffer, and R. Swanstrom, manuscript in preparation.

homologous site and in the MA/CA site. A corollary of this assumption is that the MA/CA site itself is somewhat neutral in terms of the contribution of context. We have observed that the addition of a glycine spacer (three glycines) upstream and downstream of the MA/CA site has a uniform effect on all cleavage sites in modestly reducing the rate of cleavage, suggesting that there is no significant contribution of context to a specific sequence at the MA/CA site.⁴ Thus, we interpret the similar cleavage rates of MA/CA, RT/IN, NC/SP2, TF/PR, and PR/RT in both homologous and heterologous sites as being determined largely by the cleavage site sequence itself. Hence, it should be possible to significantly increase the rate of cleavage of the slowly cleaved sequences with changes in the cleavage site sequence itself, a possibility we are exploring.

Second, the rate of cleavage is negatively regulated in the homologous site such as the CA/SP1 site, which is known to be under negative regulation in its homologous site. When placed in the MA/CA site, its rate of cleavage was consistent with removal of that negative regulation. In its homologous context, the CA/SP1 site is cleaved very slowly. Placement of this site in the MA/CA context increased its rate of cleavage to ~40% of the rate of cleavage of the MA/CA site. Similar increases in the rate of cleavage occur either when the downstream SP1/NC cleavage is blocked or when the pH is lowered (3), although in another study the enhanced rate did not result in complete cleavage at the CA/SP1 site in either both transfected cells or virus particles (6). In addition, the amino acid sequences at the CA/SP1 boundary have been predicted to adopt an α -helical structure (41); however, the NMR structure of a fragment including the C-terminal domain of capsid through the NC domain did not reveal an α -helical structure at the CA/SP1 boundary because of the flexibility of this region (42). It is possible that the amino acids at the CA/SP1 boundary region undergo a conformational change upon the initial cleavage at the SP1/NC site to function as a negative regulator on the cleavage at the CA/SP1 site, although the mechanism is still unclear.

Finally, we interpret two sites as being under local positive regulation in their homologous sites. Both the SP1/NC site and the SP2/p6 site are cleaved at significantly slower rates in the heterologous context of the MA/CA protein than in their homologous sites. Curiously, both are at the C-terminal end of a spacer peptide, and both play significant roles in the interaction with viral RNA (4, 43–45).

The SP2/p6 site is efficiently cleaved in its homologous site (3), and in the p15 (NC/SP2/p6) intermediate, its cleavage is strongly enhanced by binding of RNA to NC (4, 7). However, out of this context, *i.e.* when placed at the MA/CA site, the rate of cleavage is very slow, analogous to the rate of cleavage in the absence of RNA. It is unlikely that RNA serves as a cofactor for the protease in the cleavage of the SP2/p6 site. The alternative explanation is that RNA binding to NC promotes a conformational change that greatly enhances the presentation of the cleavage site sequence in its homologous context. It has been shown that it is the binding of RNA to NC that is important for this rate enhancement; however, the nature of the conforma-

tional change induced by RNA binding is not known. Based on this argument, the SP2/p6 site alone is suboptimal for cleavage, as seen when placed in the MA/CA site, but a local conformation is able to overcome the suboptimal sequence context to enhance recognition and/or cleavage by the viral protease when in its homologous context of NC/SP2/p6.

The SP1/NC also appears to be under positive regulation in its homologous site. This site is the first site cleaved in Gag, and its rate of cleavage in full-length Gag is significantly faster than the rate of cleavage of the MA/CA site, in contrast to its rate of cleavage when this site is placed at the MA/CA site. It should be noted that in these experiments the sequence of the NL4-3 cleavage site was used, and it contains isoleucine at the P1' position instead of methionine. The more common P1' methionine increases the rate of cleavage of the SP1/NC site several-fold when placed in this heterologous site,⁵ but it is still well below the much higher relative cleavage rate seen in the homologous site. In either form the cleavage rate of the SP1/NC site in the context of the MA/CA was far less than that in its original context, which is consistent with an earlier study demonstrated by Tritsch *et al.* (5) using *in vitro* translated Gag proteins. Cleavage at the SP1/NC site is important for the maturation of viral RNA from a low stability dimer to the high stability dimer associated with the mature virus particle (46). Although the failure of NC to bind RNA greatly reduces the rate of cleavage of the SP2/p6 site, the failure to bind RNA does not affect the generation of the p15 intermediate, suggesting that the rapid cleavage at the SP1/NC site is not dependent on NC binding to RNA. Thus, the nature of the positive regulator that makes the SP1/NC site the most rapidly cleaved site in either Gag or Gag-Pro-Pol is unknown.

With the MDR HIV-1 PR variants PR_{G2} and PR_{G4}, we observed two types of changes with respect to the relative rates of cleavage. One type of change resulted in a modest, 2–3-fold increase or decrease in the rate of cleavage relative to the MA/CA site. In our experimental design we measure rates relative to the wild type MA/CA site. If the mutant proteases had a specific change in the rate of cleavage at the MA/CA site, we would not be able to discern this effect *versus* a change at another site. Thus, it is difficult to interpret these relatively small changes with respect to mechanism. However, there was an ~10-fold relative increase by PR_{G2} in the relative rate of cleavage of the SP1/NC site, suggesting a significantly enhanced interaction between this site and the MDR protease. An examination of the number of hydrogen bonds or van der Waals interactions did not provide an obvious explanation for this change. It is possible that the mutant protease interacts with substrates in a way that at least in part compensates for the positive regulatory element that affects the rate at the homologous site. We speculate that the P1' to P4' residues of the SP1/NC site contribute to the significant proportion of the positive interactions with the PR_{G2} mutant protease (Fig. 7). The use of either SP1/NC or MA/CA half-sites also showed that neither half-site when paired with the complementary half-site was significantly active as a substrate for the wild type enzyme.

⁴ M. Potempa, S.-K. Lee, C. A. Schiffer, and R. Swanstrom, manuscript in preparation.

⁵ S.-K. Lee, M. Potempa, C. A. Schiffer, and R. Swanstrom, unpublished data.

Therefore, the chimeric substrates composed of two half-sites derived from the SP1/NC and MA/CA sites are antagonistic in their interaction with the protease, and effects across the scissile bond caused by the substitutions must alter the interaction of each half-site within the enzyme.

In this report, we demonstrate the importance of the context surrounding cleavage sites on Gag processing by using a heterologous context, the MA/CA, suggesting that interactions between the PR residues and the substrates are not confined within the active site of the protease. We suggest that context plays a different role for different sites, either being relatively neutral, providing negative regulation, or providing positive regulation. Even within a cleavage site sequence, the two half-sites can play a positive effect with each other in determining the rate of cleavage. These long range features of the regulation of cleavage rates cannot be addressed using peptide substrates and therefore point to the need to use Gag protein substrates to explore the role of context in regulating cleavage rates.

Acknowledgments—We are grateful to Nese KurtYilmaz for critically reading the manuscript. We thank Steven Lewis and Sarah Clement for assistance in generating some of the initial constructs.

REFERENCES

- Swanstrom, R., and Wills, J. W. (1997) "Synthesis, assembly, and processing of viral proteins," in *Retroviruses* (Coffin, J. M., Hughes, S. H., and Varmus, H. E., ed.), pp. 263–334, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Craven, R. C., Leure-duPree, A. E., Erdie, C. R., Wilson, C. B., and Wills, J. W. (1993) Necessity of the spacer peptide between CA and NC in the Rous sarcoma virus gag protein. *J. Virol.* **67**, 6246–6252
- Pettit, S. C., Moody, M. D., Wehbie, R. S., Kaplan, A. H., Nantermet, P. V., Klein, C. A., and Swanstrom, R. (1994) The p2 domain of human immunodeficiency virus type 1 Gag regulates sequential proteolytic processing and is required to produce fully infectious virions. *J. Virol.* **68**, 8017–8027
- Sheng, N., Pettit, S. C., Tritch, R. J., Ozturk, D. H., Rayner, M. M., Swanstrom, R., and Erickson-Viitanen, S. (1997) Determinants of the human immunodeficiency virus type 1 p15NC-RNA interaction that affect enhanced cleavage by the viral protease. *J. Virol.* **71**, 5723–5732
- Tritch, R. J., Cheng, Y. E., Yin, F. H., and Erickson-Viitanen, S. (1991) Mutagenesis of protease cleavage sites in the human immunodeficiency virus type 1 gag polyprotein. *J. Virol.* **65**, 922–930
- Wieggers, K., Rutter, G., Kottler, H., Tessmer, U., Hohenberg, H., and Kräusslich, H. G. (1998) Sequential steps in human immunodeficiency virus particle maturation revealed by alterations of individual Gag polyprotein cleavage sites. *J. Virol.* **72**, 2846–2854
- Sheng, N., and Erickson-Viitanen, S. (1994) Cleavage of p15 protein *in vitro* by human immunodeficiency virus type 1 protease is RNA dependent. *J. Virol.* **68**, 6207–6214
- Erickson-Viitanen, S., Manfredi, J., Viitanen, P., Tribe, D. E., Tritch, R., Hutchison, C. A., 3rd, Loeb, D. D., and Swanstrom, R. (1989) Cleavage of HIV-1 gag polyprotein synthesized *in vitro*: sequential cleavage by the viral protease. *AIDS Res. Hum. Retroviruses* **5**, 577–591
- Pettit, S. C., Lindquist, J. N., Kaplan, A. H., and Swanstrom, R. (2005) Processing sites in the human immunodeficiency virus type 1 (HIV-1) Gag-Pro-Pol precursor are cleaved by the viral protease at different rates. *Retrovirology* **2**, 66
- Pettit, S. C., Henderson, G. J., Schiffer, C. A., and Swanstrom, R. (2002) Replacement of the P1 amino acid of human immunodeficiency virus type 1 Gag processing sites can inhibit or enhance the rate of cleavage by the viral protease. *J. Virol.* **76**, 10226–10233
- Miller, M., Jaskólski, M., Rao, J. K., Leis, J., and Wlodawer, A. (1989) Crystal structure of a retroviral protease proves relationship to aspartic protease family. *Nature* **337**, 576–579
- Navia, M. A., Fitzgerald, P. M., McKeever, B. M., Leu, C. T., Heimbach, J. C., Herber, W. K., Sigal, I. S., Darke, P. L., and Springer, J. P. (1989) Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. *Nature* **337**, 615–620
- Pearl, L. H., and Taylor, W. R. (1987) A structural model for the retroviral proteases. *Nature* **329**, 351–354
- Pettit, S. C., Simsic, J., Loeb, D. D., Everitt, L., Hutchison, C. A., 3rd, and Swanstrom, R. (1991) Analysis of retroviral protease cleavage sites reveals two types of cleavage sites and the structural requirements of the P1 amino acid. *J. Biol. Chem.* **266**, 14539–14547
- Prabu-Jeyabalan, M., Nalivaika, E., and Schiffer, C. A. (2002) Substrate shape determines specificity of recognition for HIV-1 protease. Analysis of crystal structures of six substrate complexes. *Structure* **10**, 369–381
- Ozen, A., Halilolu, T., and Schiffer, C. A. (2011) Dynamics of preferential substrate recognition in HIV-1 protease. Redefining the substrate envelope. *J. Mol. Biol.* **410**, 726–744
- Billich, S., Knoop, M. T., Hansen, J., Strop, P., Sedlacek, J., Mertz, R., and Moelling, K. (1988) Synthetic peptides as substrates and inhibitors of human immunodeficiency virus-1 protease. *J. Biol. Chem.* **263**, 17905–17908
- Darke, P. L., Nutt, R. F., Brady, S. F., Garsky, V. M., Ciccarone, T. M., Leu, C. T., Lumma, P. K., Freidinger, R. M., Veber, D. F., and Sigal, I. S. (1988) HIV-1 protease specificity of peptide cleavage is sufficient for processing of gag and pol polyproteins. *Biochem. Biophys. Res. Commun.* **156**, 297–303
- Ermoloeff, J., Lin, X., and Tang, J. (1997) Kinetic properties of saquinavir-resistant mutants of human immunodeficiency virus type 1 protease and their implications in drug resistance *in vivo*. *Biochemistry* **36**, 12364–12370
- Kräusslich, H. G., Ingraham, R. H., Skoog, M. T., Wimmer, E., Pallai, P. V., and Carter, C. A. (1989) Activity of purified biosynthetic proteinase of human immunodeficiency virus on natural substrates and synthetic peptides. *Proc. Natl. Acad. Sci. U.S.A.* **86**, 807–811
- Schock, H. B., Garsky, V. M., and Kuo, L. C. (1996) Mutational anatomy of an HIV-1 protease variant conferring cross-resistance to protease inhibitors in clinical trials. Compensatory modulations of binding and activity. *J. Biol. Chem.* **271**, 31957–31963
- Tözsér, J., Bláha, I., Copeland, T. D., Wondrak, E. M., and Oroszlan, S. (1991) Comparison of the HIV-1 and HIV-2 proteinases using oligopeptide substrates representing cleavage sites in Gag and Gag-Pol polyproteins. *FEBS Lett.* **281**, 77–80
- Huang, L., Li, Y., and Chen, C. (2011) Flexible catalytic site conformations implicated in modulation of HIV-1 protease autoprocessing reactions. *Retrovirology* **8**, 79
- Adams, S. R., Campbell, R. E., Gross, L. A., Martin, B. R., Walkup, G. K., Yao, Y., Llopis, J., and Tsien, R. Y. (2002) New biarsenical ligands and tetracysteine motifs for protein labeling *in vitro* and *in vivo*. Synthesis and biological applications. *J. Am. Chem. Soc.* **124**, 6063–6076
- Griffin, B. A., Adams, S. R., Jones, J., and Tsien, R. Y. (2000) Fluorescent labeling of recombinant proteins in living cells with FAsH. *Methods Enzymol.* **327**, 565–578
- Griffin, B. A., Adams, S. R., and Tsien, R. Y. (1998) Specific covalent labeling of recombinant protein molecules inside live cells. *Science* **281**, 269–272
- Adachi, A., Gendelman, H. E., Koenig, S., Folks, T., Willey, R., Rabson, A., and Martin, M. A. (1986) Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. *J. Virol.* **59**, 284–291
- Lee, S. K., and Hacker, D. L. (2001) *In vitro* analysis of an RNA binding site within the N-terminal 30 amino acids of the southern cowpea mosaic virus coat protein. *Virology* **286**, 317–327
- Studier, F. W., and Moffatt, B. A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189**, 113–130
- Tomasselli, A. G., Olsen, M. K., Hui, J. O., Staples, D. J., Sawyer, T. K., Heinrikson, R. L., and Tomich, C. S. (1990) Substrate analogue inhibition and active site titration of purified recombinant HIV-1 protease. *Biochemistry* **29**, 264–269

31. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685
32. King, N. M., Melnick, L., Prabu-Jeyabalan, M., Nalivaika, E. A., Yang, S. S., Gao, Y., Nie, X., Zepp, C., Heefner, D. L., and Schiffer, C. A. (2002) Lack of synergy for inhibitors targeting a multi-drug-resistant HIV-1 protease. *Protein Sci.* **11**, 418–429
33. Blommel, P. G., and Fox, B. G. (2005) Fluorescence anisotropy assay for proteolysis of specifically labeled fusion proteins. *Anal. Biochem.* **336**, 75–86
34. Campbell, E. M., Perez, O., Anderson, J. L., and Hope, T. J. (2008) Visualization of a proteasome-independent intermediate during restriction of HIV-1 by rhesus TRIM5 α . *J. Cell Biol.* **180**, 549–561
35. Jolley, M. E. (1996) Fluorescence polarization assays for the detection of proteases and their inhibitors. *J. Biomol. Screen.* **1**, 33–38
36. Dauber, D. S., Ziermann, R., Parkin, N., Maly, D. J., Mahrus, S., Harris, J. L., Ellman, J. A., Petropoulos, C., and Craik, C. S. (2002) Altered substrate specificity of drug-resistant human immunodeficiency virus type 1 protease. *J. Virol.* **76**, 1359–1368
37. Liu, Z., Wang, Y., Brunzelle, J., Kovari, I. A., and Kovari, L. C. (2011) Nine crystal structures determine the substrate envelope of the MDR HIV-1 protease. *Protein J.* **30**, 173–183
38. Richards, A. D., Roberts, R., Dunn, B. M., Graves, M. C., and Kay, J. (1989) Effective blocking of HIV-1 proteinase activity by characteristic inhibitors of aspartic proteinases. *FEBS Lett.* **247**, 113–117
39. Hyland, L. J., Tomaszek, T. A., Jr., and Meek, T. D. (1991) Human immunodeficiency virus-1 protease. 2. Use of pH rate studies and solvent kinetic isotope effects to elucidate details of chemical mechanism. *Biochemistry* **30**, 8454–8463
40. Tomaszek, T. A., Jr., Moore, M. L., Strickler, J. E., Sanchez, R. L., Dixon, J. S., Metcalf, B. W., Hassell, A., Dreyer, G. B., Brooks, L., and Debouck, C. (1992) Proteolysis of an active site peptide of lactate dehydrogenase by human immunodeficiency virus type 1 protease. *Biochemistry* **31**, 10153–10168
41. Accola, M. A., Höglund, S., and Göttlinger, H. G. (1998) A putative α -helical structure which overlaps the capsid-p2 boundary in the human immunodeficiency virus type 1 Gag precursor is crucial for viral particle assembly. *J. Virol.* **72**, 2072–2078
42. Newman, J. L., Butcher, E. W., Patel, D. T., Mikhaylenko, Y., and Summers, M. F. (2004) Flexibility in the P2 domain of the HIV-1 Gag polyprotein. *Protein Sci.* **13**, 2101–2107
43. Kaye, J. F., and Lever, A. M. (1998) Nonreciprocal packaging of human immunodeficiency virus type 1 and type 2 RNA. A possible role for the p2 domain of Gag in RNA encapsidation. *J. Virol.* **72**, 5877–5885
44. Morellet, N., Druillennec, S., Lenoir, C., Bouaziz, S., and Roques, B. P. (2005) Helical structure determined by NMR of the HIV-1 (345–392)Gag sequence, surrounding p2. Implications for particle assembly and RNA packaging. *Protein Sci.* **14**, 375–386
45. Roldan, A., Russell, R. S., Marchand, B., Götte, M., Liang, C., and Wainberg, M. A. (2004) *In vitro* identification and characterization of an early complex linking HIV-1 genomic RNA recognition and Pr55Gag multimerization. *J. Biol. Chem.* **279**, 39886–39894
46. Hill, M. K., Shehu-Xhilaga, M., Crowe, S. M., and Mak, J. (2002) Proline residues within spacer peptide p1 are important for human immunodeficiency virus type 1 infectivity, protein processing, and genomic RNA dimer stability. *J. Virol.* **76**, 11245–11253